

1-1-2003

Oil and Water: Why Retribution and Repentance Do Not Mix

Sherry F. Colb

Cornell Law School, sfc44@cornell.edu

Follow this and additional works at: <http://scholarship.law.cornell.edu/facpub>



Part of the [Criminal Law Commons](#)

Recommended Citation

Colb, Sherry F., "Oil and Water: Why Retribution and Repentance Do Not Mix" (2003). *Cornell Law Faculty Publications*. Paper 625.
<http://scholarship.law.cornell.edu/facpub/625>

This Article is brought to you for free and open access by the Faculty Scholarship at Scholarship@Cornell Law: A Digital Repository. It has been accepted for inclusion in Cornell Law Faculty Publications by an authorized administrator of Scholarship@Cornell Law: A Digital Repository. For more information, please contact jmp8@cornell.edu.

OIL AND WATER: WHY RETRIBUTION AND REPENTANCE DO NOT MIX

*By Sherry F. Colb**

Professor Herbert Morris' work combines a strong affinity for justice with an equally powerful impulse toward humane and compassionate responses to transgression. He speaks of the retributive value of punishment as an expression not only of the need to right a wrong but also as an essential component of integrating the transgressing member of society back into the fold. This is an appealing vision at a time when public policy makers take retributive justice to mean pure vengeance.

In elaborating on the importance of humane and respectful punishment, Morris provides a useful analogy to a parent teaching a child right from wrong. Unlike the vengeance of the criminal justice system we have in place today, a system that lives up to few ideal visions of the retributive state, the model of the parent disciplining a child reflects an essential role for teaching and nurturing in punishment.

The analogy invites us to view the wrongdoer from the perspective of someone who not only feels aggrieved by what he has done but also cares about him. His act is grave enough that doing nothing retributive would leave a rift between him and society. The objective of punishment, however, is not exclusively to cause suffering but to convey a message that cannot otherwise be communicated. Saying "what you have done is wrong," coupled with punishment, drives home just how seriously we, the parent-figure, find the misconduct of the perpetrator.

In this paper, I will argue that this analogy helps underscore why punishment cannot effectively transmit values from society to offender.

* Professor of Law and Judge Frederick B. Lacey Scholar, Rutgers Law School-Newark. The author thanks Michael C. Dorf and the participants in the Quinnipiac University School of Law Conference on the Work of Herbert Morris for reading earlier drafts and providing helpful commentary and suggestions for improving the work. The author also gratefully acknowledges Justin Petruzzelli, Tim Madden, and Jung Kim for their invaluable research assistance.

In making this argument, I will draw on three sources in various permutations: my own intuitions about the subjective experience of being punished, an interview I conducted with former child molester Leroy Hendricks, and empirical studies of intrinsic motivation in the field of social psychology.

I. THE PARENT-CHILD ANALOGY

Since a parent punishes a child, at least in part, to turn him into a better person, one could describe such punishment as utilitarian or consequentialist. The idea is that I, the parent, will cause suffering to my child right now to prevent her from acting this way in the future and to encourage her to conform to my vision of a better person. One could describe this particular strategy as operant conditioning: reinforcing the internal inclination to do right and extinguishing the inclination to do wrong. However, this is notably *not* what Morris means when he says that society should discipline its transgressors in the way that a parent disciplines a child.

I suggest that Morris would disapprove of the operant conditioning approach to parental and societal punishment, because he specifically eschews the notion of giving a transgressor a pill that would turn him into a better person.¹ If the goal were simply to make him better while minimizing overall suffering, such a pill would be ideal. Indeed, a parent whose child was cruel and nasty might jump at the opportunity to administer a pill that would automatically change the child into a kind and thoughtful person.

Surely, there are consequential risks attached to permitting the use of such “good person” pills. Quite apart from the risks, however, Morris would object to the idea in principle. To medicate a person in an

1. See Herbert Morris, *A Paternalistic Theory of Punishment*, 18 AM. PHIL. Q. 263, 265-66 (1981) [hereinafter Morris, *A Paternalistic Theory*].

[U]nacceptable to this theory [of legitimate punishment] would be any response [to wrongdoing] that sought the good of a wrongdoer in a manner that bypassed the human capacity for reflection, understanding, and revision of attitude that may result from such efforts. Any punitive response to a fully responsible being, then, and it might be no more than the giving of an evil-tasting pill or some form of conditioning, that directly in some causal way, with or without the agent's consent, sought to bring about a good, say, instantaneous truth or aversion to acting violently, would be incompatible with this constraint.

Id. Herbert Morris, *The Decline of Guilt*, 99 ETHICS 62, 71 (1988) (criticizing “tendencies toward viewing antisocial conduct as pathological, a matter for therapy, not punishment”).

effort to change her from bad to good does not, in Morris' scheme, treat a person with respect as an autonomous individual. In deontological terms, it treats her as a means rather than an end in herself. We treat the person as an end in herself precisely by using an unpleasant (but not brainwashing) experience to communicate openly just how offensive and normatively unacceptable her behavior is.

Imagine the following example: my son laughs at a person who is mentally retarded. I notice this laughter and find myself angry with my son and distressed that a person with whom I identify could be so cruel and unfeeling. I call my son over and tell him he will go without television (or the Internet or the telephone) for a week as punishment for his cruelty. My goal is not to transform him by exertion of power, but rather, by Morris' analysis, to make my son aware, in a very tangible and concrete way, that his behavior has caused a rift between us, one that can only be rectified by his serving "time." He understands through his experience of punishment, as he might not were I to use only words, that his behavior has violated my set of norms.² On this account, punishment is a necessary, metaphysically transformative experience for my son and a cleansing experience for our relationship.

Morris contends that punishment for criminals should similarly convey a message about how wrongfully they have acted. Their wrongful behavior, in addition to causing pain, creates a distance between them and the surrounding society. To ignore what they have done—or simply to express disapproval without punishment—would not fully accomplish the reintegration that must occur at a moral, metaphysical level before the wrongdoer can regain his place within the society he has injured.

This approach is theoretically quite appealing. It respects the

2. Others have developed related expressivist theories of punishment. See, e.g., Dan M. Kahan, *What Do Alternative Sanctions Mean?*, 63 U. CHI. L. REV. 591 (1996); Richard H. Pildes & Elizabeth S. Anderson, *Slings Arrows at Democracy: Social Choice Theory, Value Pluralism, and Democratic Politics*, 90 COLUM. L. REV. 2121 (1990); Richard H. Pildes & Cass R. Sunstein, *Reinventing the Regulatory State*, 62 U. CHI. L. REV. 1 (1995). Such theories have been criticized for various reasons, including the potential alternative ways of communicating disapproval without having to inflict suffering. See Matthew D. Adler, *Expressive Theories of Law: A Skeptical Overview*, 148 U. PA. L. REV. 1363 (2000); Matthew D. Adler, *Linguistic Meaning, Nonlinguistic "Expression," and the Multiple Variants of Expressivism: A Reply to Professors Anderson and Pildes*, 148 U. PA. L. REV. 1577 (2000). I do not here address the question of whether expressivist theories of punishment are persuasive as a general matter. My focus is instead on the separate but related issue of whether punishment is in fact incompatible with the objective of truly "reaching" the wrongdoer in a communicative sense.

offender as a moral agent to be taken seriously. Yet, my own hardly unique experience of suffering punishment as a child raises doubts. To generalize, when a child is punished, she feels wronged. In my case, this was so even when, in theory, I agreed with the motivation for the punishment—for example, one should not yell at one's elders. I felt rage at the authority figure who disciplined me, and I resented the punishment. "How dare my mother judge me," I thought, "when I had a good reason for doing what I did? I shall sit through my punishment, but I do not accept the judgment that goes with it."

My parent or teacher could not take away my belief in the rectitude of my actions, even if, with her superior strength, she could take away privileges or otherwise inflict suffering. Instead, I experienced punishment most often as alienating, arbitrary, and wrongfully injurious.

Perhaps this reaction to punishment characterizes only children, and perhaps I was not a typical child. Think now of prison, a far cry from an ideal retributive environment, but a useful place to consider in evaluating what punishment can and cannot do. Popular fiction illustrates a point well known to those with substantial experience in dealing with prisoners. In *The Shawshank Redemption*, for example, an old prisoner played by actor Morgan Freeman, Ellis Boyd "Red" Redding, speaks to a younger prisoner, Andy Dufresne.³ The latter, played by Tim Robbins, has been wrongly accused and convicted of a crime. Redding responds to Dufresne's protestations of innocence by remarking that of all the people at Shawshank prison, he (Redding) appears to be the only one who is actually guilty of committing a crime.

As high recidivism rates confirm, our bursting prisons do not occasion much true repentance,⁴ even though some prisoners may acknowledge that they have violated the law. While admitting the acts for which they were convicted, many have excuses and explanations for why their deeds were not as bad as everyone says.

3. THE SHAWSHANK REDEMPTION (Columbia Pictures 1994).

4. See Terry Morgan, *HITS/SMART: Washington State's Crime-Fighting Tool*, FBI LAW ENFORCEMENT BULL., Vol. 71, No. 2, 1-10, 2 (2002).

[I]n 1997, only 44 percent of persons under state supervision successfully completed their term of supervision compared with 70 percent in 1984. Moreover, police arrested a high percentage of these individuals for new felonies. Over 10 years earlier, a study in Los Angeles County concluded that recidivism rates for high-risk offenders ranged between 50 and 75 percent, while a Philadelphia study conducted nearly 30 years ago found that 7 percent of the criminal population committed approximately 70 percent of all violent crime.

Id.

In addition, prisoners often develop norms of behavior that are different from those of the surrounding culture. I understand them to include a hierarchy of crimes in which sexual offenses against children are considered especially heinous and ratting on one's friends is entirely unforgivable. Some crimes, by contrast, like a revenge killing of a person perceived as a wrongdoer, might earn respect and admiration, rather than condemnation, from other prisoners. We engage in a futile attempt to integrate a person back into the law-abiding community when we surround him with others who reject that community's values.

Success in finding an approach that will "reach" those people who have committed wrongful acts will turn largely on an accurate understanding of human nature. Even though one's approach might not be "consequentialist," it is nonetheless necessary to consider the actual effects of societal condemnation if the hope is to reach and communicate norms to living individuals in the real world. If a person punished for wrongdoing does not comprehend that his suffering is a result of his actions, the punishment is unlikely to have any communicative significance for him.

For example, setting fire to a person's home as retaliation for his cheating the arsonist in business gets the "message" across only if the arson victim understands the connection between what he has done and what he suffers. To communicate, we must speak in a language that can be readily understood by our intended audience. This is one reason for the insanity defense—the meaninglessness of penalties for the truly insane.⁵

Consider an example that arises occasionally in daily life. Some stranger insults you in a public place. You try to console yourself by imagining that the obnoxious person has a miserable life. The response is not wholly satisfying, however, because you want the transgressor to be miserable *because* of her transgression against you so that she links the two and sees how she has brought the suffering on herself by being bad.⁶

5. Cf. *Ford v. Wainwright*, 477 U.S. 399, 409 (1986) (prohibiting execution of the insane and stating that "[W]e may seriously question the retributive value of executing a person who has no comprehension of why he has been singled out and stripped of his fundamental right to life.") (citations omitted).

6. Matthew D. Adler, *Expressive Theories of Law: A Skeptical Overview*, 148 U. PA. L. REV. 1363, 1421 (2000) ("'Poetic justice' involves the wrongdoer's undergoing a consequence that appropriately could be visited upon him in retribution but which was not produced in that way . . .") (citing ROBERT NOZICK, *PHILOSOPHICAL EXPLANATIONS* 363, 369-70 (1981)).

Human nature, I believe, leads many people to react defensively to punishment that they *know* is a response to their perceived wrongdoing, however. Under circumstances of condemnation, people seem to lack the capacity to understand that what they did really was wrong. Children react this way to parental punishment (or at least I did), and prisoners react this way to their sentences. If, as I suspect, this is because people receive punishment as a blow to their egos and thus feel anger rather than contrition in response, then punishment usually will fail at communicating to a transgressor the seriousness of his transgression. His tendency to be defensive, in other words, rather than repentant, precludes his truly hearing and accepting the message society tries to convey.

In keeping with the view that punishment is the proper approach to healing the rift a transgressor creates between himself and society, Herbert Morris rejects what one might call the “therapeutic state.”⁷ The therapeutic state is a government that exercises social control over harmful and destructive behavior as a pathology to be treated, rather than an evil to be punished. In the therapeutic state, for example, if I were to steal property from a neighbor, I might be confined in an institution so long as I posed a threat of further thefts. Consistent with the goal of my confinement, I might be given a regimen of medicines and therapies to reduce my inclination toward theft in the future. Once I changed and no longer posed a danger to those around me, I could then reemerge into society and resume my life.

Morris argues that this regime is insulting to responsible human beings. For people who are insane and thus unable to regulate their behavior or understand the nature or morality of what they do, a system of therapeutic intervention is often the only recourse. It is cruel and meaningless to punish a person who cannot exercise autonomy or has no idea what he is doing, much as it would be cruel and meaningless to punish a person for a spasm that causes harm. For the overwhelming majority of the population, however, the government should permit

7. See Morris, *A Paternalistic Theory*, *supra* note 1, at 269. Morris explained that his:

theory would regard as morally unacceptable a response, conditioning or otherwise, that had as its goal, not just aversion to doing wrong, but obliteration of one's capacity to choose to do so. What must be aimed at is that the afflicted become autonomous not automatons. There must be freedom to disobey, for the moral price is too high that is paid in purchasing immunity from temptation and guaranteed conformity.

Id.

individuals to retain their freedom and autonomy unless and until they do something wrong. At that point, the government should punish them for their deeds and permit eventual reintegration.

The debate about the therapeutic state has great practical significance with the emergence of “civil” commitment of sexually violent predators after their prison sentences have run.⁸ Such predators are people who, under conventional definitions of criminal responsibility, were responsible for the harm and suffering they caused. They made the decision to seduce and molest children, for example, because they felt a desire and an impulse toward those children, and they acted on that impulse despite the fact that they fully understood that such conduct violated the law.

We know that the predators understood the illegality of their behavior because, in most instances, they went to great pains to conceal their conduct, including warnings and threats to their victims not to tell anyone what happened. Since in many cases, moreover, the predators served prison sentences for their behavior, we know that the law considered them “responsible” enough to qualify for punishment. Yet, once their sentences ran and they were poised to reenter society without having decisively changed into better people, states reacted by legislating a form of civil commitment that would treat them as pathological, or “mad,” rather than “bad.”⁹

Stephen Morse considers this trend quite alarming and agrees with Herbert Morris’ views in this regard.¹⁰ Morse, like Morris, argues that only nonresponsible people should be subject to civil “therapeutic” commitment and that a responsible person who is dangerous because of the evil we predict he will do—whether or not due in part to a mental illness—should be left alone, unless and until he acts on his inclinations and renders himself subject to our penal system. Both Morris and Morse believe it is an assault on human dignity and autonomy to do

8. See, e.g., Kansas Sexually Violent Predator Act, KAN. STAT. ANN. § 59-29a01 (2001) (“Because the existing civil commitment procedures . . . are inadequate to address the special needs of sexually violent predators and the risks they present to society, the legislature determines that a separate involuntary civil commitment process for the potentially long-term control, care and treatment of sexually violent predators is necessary.”).

9. See, e.g., MO. REV. STAT. §§ 632.480 to 632.513 (1998) (current version at MO. REV. STAT. §§ 632.480-632.513 (2002)); N.J. STAT. ANN. §§ 30:4-27.1 to 30:4-27.38 (1998); WASH. REV. CODE §§ 71.09.010 to 71.09.800 (1998) (current version at WASH. REV. CODE §§ 71.09.010 to 71.09.800 (2002)).

10. See Stephen J. Morse, Essay, *Uncontrollable Urges and Irrational People*, 88 VA. L. REV. 1025 (2002).

otherwise.

But what if human beings are simply not programmed to reflect on their misdeeds in response to punishment? If one could predict danger with a high level of accuracy, and if the necessity for continued confinement of a patient were truly revisited regularly, Morse might accept civil confinement of such people to protect the public from an ongoing danger.¹¹ Morris, however, would not urge this course except for the insane. For all others, punishment, and not therapeutic commitment, is the appropriate response.

II. PUNISHMENT AS FUTILE: INTERVIEW WITH LEROY HENDRICKS

I have argued that punishment is inherently bound to fail at teaching offenders right from wrong. People respond to punishment with exculpatory protestations, anger, and a childlike belief that it is they who are right and the punishing body that is wrong. Dreams of revenge on testimonial witnesses, judges, and prosecutors surface as responses to the experience of punishment.¹²

In voicing the concern that punishment is, as a factual matter, an ineffective way to convey values to offenders, I think of Leroy Hendricks, the first person to be confined pursuant to the Kansas Sexually Violent Predator Act ("SVPA").¹³ After Hendricks served a ten-year prison sentence for child molestation, Kansas held a hearing under its statute and found, beyond a reasonable doubt, that Hendricks was mentally abnormal and dangerous. After the hearing, Hendricks was placed in a facility in Larned, Kansas, which was housed within a prison but separated and defined by the distinct agenda of civilly confining sexually violent predators.

Hendricks challenged his confinement as unconstitutional on a variety of grounds. The essence of his primary claim arose from the view that civil confinement for sexual predators is criminal punishment and should therefore carry the same constitutional limits as would any

11. *Id.* at 1026 n.5 ("I agree that many criminal justice practices impose harsh and disproportionate punishments and that perhaps pure preventive detention would be justified by public safety concerns, if predictive accuracy were sufficiently great.").

12. Punishment may nonetheless work as an effective general deterrent, if those watching the punishment process learn by modeling behavior they see rewarded or punished. They might not feel the same investment in behavior at which they have not themselves been caught or for which they have not been punished.

13. KAN. STAT. ANN. §§ 59-29a01 to 59-29a15 (1994) (current version at KAN. STAT. ANN. §§ 59-29a01 to 59-29a20 (2002)).

other criminal penalty. The U.S. Supreme Court rejected Hendricks' claim and held that the confinement of sexually violent predators is a form of civil commitment rather than a criminal sanction and therefore need not conform to constitutional mandates specific to the criminal law.¹⁴

Morris would likely sympathize with Hendricks' predicament. As a sane person who could be, and indeed was, held criminally responsible for his actions, Hendricks does not fall within the very small class of individuals whose antisocial behavior ought to be "therapeutically" managed. Hendricks, in other words, had already been adjudicated "bad" rather than "mad" and should have been punished rather than "treated." While Hendricks believed that he *was* being illegally punished twice for the same offense, Morris might say that if Hendricks was indeed being managed therapeutically, as Kansas claimed, such a response to antisocial conduct is entirely inappropriate for people who, like Hendricks, have the capacity to take responsibility for their actions.

In some ways, perhaps, the distinction here is subtle. Both Morris and Hendricks' attorneys might wish to invalidate the Kansas practice of confining people in a hospital for violent predators. The distinction is important, though, for reasons that became increasingly clear to me after I had the opportunity to talk with Leroy Hendricks. On a theoretical level, based on the arguments he actually made before the Supreme Court, if Hendricks had not already been punished for his acts of child molestation, he might well have had little to complain about vis-à-vis his confinement as a predator.

If confinement in an institution is "punishment," then Hendricks earned that punishment by committing a crime. He was found beyond a reasonable doubt to have done so. This was the predicate for the conclusion that he would likely do so again. He might, of course, have challenged the prospective rather than retrospective nature of the commitment inquiry, but—if he was right that what was really happening was punishment—then this "not punishment" approach was

14. *Kansas v. Hendricks*, 521 U.S. 346, 361 (1997). *See also* *Kansas v. Crane*, 534 U.S. 407, 412 (2002). The Court also rejected a second argument that Hendricks made regarding the overbreadth of civil commitment for the "mentally abnormal" and dangerous, rather than limited just to the mentally ill. *Hendricks*, 521 U.S. at 358-59. *See* Sherry F. Colb, *Insane Fear: The Discriminatory Category of "Mentally Ill and Dangerous"*, 25 NEW ENG. J. ON CRIM. & CIV. CONFINEMENT 341, 364-66 (1999) (critiquing the distinction between the mentally ill and everyone else and thereby rejecting the second argument advanced by Hendricks and rejected by the Supreme Court).

pretextual and, accordingly, a small hurdle for the state to overcome.

Morris' complaint about the SVPA, by contrast, would be a deeper one. He would say that Leroy Hendricks is a criminal who is not insane and therefore—regardless of whether he was or was not punished in the past—should not be subject to a treatment-based approach. Instead, Hendricks should be punished, if he was not already punished, and that experience of punishment should be calibrated to convey to him the seriousness of what he has done in a way that he can comprehend as an autonomous individual and to which he will perhaps respond by bettering himself.

In December of 1998, I visited Leroy Hendricks in the Larned, Kansas facility. Hendricks and I spoke for several hours about how he came to begin his career of sexual relations with minors and children, what he thought about his present circumstances, and how he spent his days in Larned. The conversation was very illuminating. He vehemently referred to his confinement as “punishment” and “incarceration.”¹⁵

Hendricks explained that each time he was caught molesting children, he pleaded guilty, rather than putting the government to the expense of a trial. In return for his cooperation, he received shorter sentences for his crimes than he might otherwise have received. In connection with the specific crime for which he had most recently been incarcerated, for example, he could have spent a total of twenty years in prison. Because he pleaded guilty, however, he was required to spend only ten years behind bars.

He played the game, he asserted, and then the authorities changed the rules. He indicated that in the future, he would never agree to plead guilty again. His attitude was, let them waste their time trying him. Why should he do them any favors? He had apparently lost his trust in the government as an honorable competitor in the game of criminal litigation.

There is something quite compelling about Hendricks' arguments. If one is making a deal with the government, it should not be so simple for the government—in the guise of doing something “civil” rather than criminal—to avoid giving the person the benefit of his bargain. In the language of the federal sentencing guidelines, Hendricks was entitled to a reduction in his sentence for “acceptance of responsibility.” He accepted responsibility by pleading guilty, after all.

15. Interview with Leroy Hendricks in Larned, Kansas (Dec. 23, 1998) [hereinafter Leroy Hendricks Interview].

Consider, however, how Hendricks thought about punishment. Even though he “accepted responsibility,” he did not perceive his “acceptance” as embodying any sense of guilt or remorse. Instead, it was more of a strategic move in a card game against the government. This attitude does not seem likely to coincide with the contrition and repentance that we might hope would be a component of accepting responsibility.

Hendricks was like a child being punished for his bad behavior. He was ready to “accept” the punishment but does not accept the proposition that he truly deserved that punishment. He was therefore extremely unlikely to undergo a moral transformation through the communicative mechanism that Morris posits.

Hendricks described how he spends his time in the program and what some of the contrasts were between this and his prior experience as a prisoner. He said he liked the people who worked in the Larned Sexually Violent Predator Unit a great deal. He believed they were honestly committed to trying to help him become someone who was safe to be on the outside.

I asked Hendricks whether he thought that maybe someday he would no longer be sexually attracted to children. He said he doubted that would happen and could not see how his attraction would change. He compared his predicament with that of a person who wants to quit eating hamburgers because they are bad for his health. You are not going to get the person to stop feeling a desire for those hamburgers, he explained. All you can do is train him to resist the temptation as best he can. Hendricks seemed to view his deviant sexual desires as comparable to the compulsions that an addict, an alcoholic, or a binge-eater feels. His aim was to become a nonpracticing or “recovering” pedophile.

I asked Hendricks how this recovery would happen. He said he had been receiving cognitive therapy, which has been quite successful in the treatment of clinical depression. For depressives, the therapy involves a patient’s articulating false beliefs that she holds and coming up with more accurate substitutes for those beliefs. An example is the depressive’s common belief that no one likes her and that she fails at everything. Whenever these thoughts pop into her head, the therapist instructs her to write them down and then write a response to them, such as, “it is true that some people don’t like me but others do, like my neighbor Tom and my coworker Jill; they like me a lot and choose to spend a great deal of time with me,” or “I don’t always succeed at

everything I try, and failing is painful, but I have often succeeded at my endeavors. I graduated from high school, for example, and I successfully interviewed for the job I now hold. Even when I don't get what I want, sometimes just trying something I was afraid to try is a success in and of itself." Though seemingly quite basic (and perhaps childish), this approach has worked at bringing dramatic improvement—without medication—even to suicidally depressed patients.¹⁶

Although I was familiar with the literature on cognitive therapy for depression, I had a difficult time imagining what cognitive therapy for pedophiles would look like, and I said so to Hendricks. He responded that he used to have a lot of false beliefs that facilitated his acts of child molestation. Examples included his belief that the children with whom he had sexual relations were actually enjoying themselves. Sex is fun, he said, and he never forced himself on anyone. This became a major theme for Hendricks in our conversation—the distinction between the rapist who was also part of the Larned program (and whom Hendricks subtly pointed out to me later when we ate lunch with the whole group) and himself, who seduced rather than raped his victims.

Hendricks explained that he didn't mean to hurt his victims but did not formerly believe that he was in fact hurting them. Instead, he thought that society simply disapproved of his ways but had no good reason other than prudishness for that disapproval. In addition to denying the suffering he caused his victims, he implied that he had also denied the value of those people as individuals: "[W]e used 'em, throw 'em away, put 'em in the trash can, and walked away from them. Nothing doing. That's it."¹⁷

As part of his cognitive therapy, Hendricks considered how people victimized by child molestation feel; he took their perspective on what had happened to them. He was surprised by what he learned, both because he had thought his acts were fun for everyone, and because he had denied in his own mind that the victims even had a perspective. He

16. See, e.g., AARON T. BECK, COGNITIVE THERAPY AND THE EMOTIONAL DISORDERS 301 (1976) (citations omitted) ("These patients, although not substantially helped by drug therapy, showed prompt and sustained improvement with cognitive-behavioral therapy as reflected by their scores on the Hamilton and Beck depression measures . . ."). "Patients in the psychotherapy group showed more rapid improvement in the degree of hopelessness than did the drug-treated patients. Because hopelessness is an index of suicide risk, it is possible that this psychotherapy was more effective than chemotherapy in shortening the period of suicide risk." *Id.* at 304.

17. Leroy Hendricks Interview, *supra* note 15.

began responding internally to his ideas about victims, and this motivated him to want to stop his molesting behavior.

I asked Hendricks whether he would have liked, in the first instance, to have been sentenced to the Larned program instead of prison. He said that although he was locked up, he definitely preferred the Larned program to prison. He believed he had wasted all of those years in the penitentiary. Moreover, he had apparently never really internalized the importance of not molesting children.

I inquired of Hendricks whether he was close to anyone outside of the institution. It turned out that he had been married twice to the same woman but that neither marriage had worked out. He continued to be close with his mother, as he had always been. Now in her eighties, he told me, his mother had disclosed something to him on a visit to Larned that he had never known before. As a child, his mother was molested. Presumably, she had not told Hendricks about that experience previously, because she thought that as a child molester himself, he might not be very sympathetic. Though she had loved him through all of his convictions and prison sentences, she did not trust him with her painful past.

Hendricks' mother finally decided to reveal her secret, however, on one of her visits to Larned. Perhaps it was because she saw a change in her son. She might have observed some remorse for what he had done to the many children he molested. For whatever reason, she decided that she could share with him the story of what had happened to her. His reaction to that disclosure was to become very angry about her victimization and to connect her suffering with that of his own victims. He felt ashamed of what he had done and felt he had been part of a group of bad people, one of whom had caused devastating pain to his own mother.

I was struck by the contrast between Hendricks' feelings, on the one hand, about the Kansas government and its decision to place him in the Larned facility, and on the other, about the reality of the hospital setting itself. Though he had a few complaints about some of the personnel at Larned (Hendricks confided that one, for example, apparently thought that all women were good and that all men were bad), he felt that they were, on the whole, committed to helping him and they had already made him want to change and given him faith that maybe he could. He said that his odds of re-offending, in his mind, had

dropped quite a bit since he had first entered the program.¹⁸

I asked Hendricks about the interactions, if any, between him and the prisoners housed in the same building. He indicated they did not have very much contact, other than in the cafeteria, but when they did, the prisoners mocked him and his fellow patients. The prisoners believed, according to him, that patients fell beneath prisoners in terms of their status. I asked him for his reaction to this attitude. He responded that he disagreed with the view of the prisoners. He pointed out that he was able to wear his own clothing and play guitar and was not, in fact, a prisoner serving time for having committed a crime. He seemed to think the prisoners were foolish to consider themselves above him. “[Y]ou think we’re sex offenders? Ninety-nine percent of them are, too, worse than we are. And they’ll be over here before long.”¹⁹

Hendricks accordingly expressed disappointment in a company representative from whom he had tried to order model airplane materials over the telephone. The woman on the phone had apparently become unfriendly when Hendricks explained that she was to send the materials to him at the Sexually Violent Predator Unit. She asked whether this was some sort of a prison, and he explained—with some pride—that he was not a prisoner but, rather, was civilly committed for treatment. When this revelation did not alter her negative tone, Hendricks decided to take his business elsewhere and felt that the woman to whom he had spoken was prejudiced and was missing out on a good customer. His attitude suggested a level of self-esteem and confidence in his current status that seemed both surprising and very positive.

In elaborating strategies for reducing the odds of his re-offending in the future, Hendricks spoke of distractions that would place some distance between him and the source of temptation. He explained that it was critical that he protect himself from tempting situations because that was the most effective way to avoid harming anyone. He posed a hypothetical example in which he would be in a convenience store making a purchase and find himself near a young child to whom he was attracted. In such a circumstance, he would walk away from the child and perhaps even leave the store. There would be children around him once he left the institution, he explained, and he would have to devise ways of handling the impulses that would inevitably arise.

Hendricks never enjoyed prison and seemed to prefer the

18. *Id.* (“Back then, I’d give myself about a five. Today I’d give myself a two. [Back then, there] would’ve been a 50/50 chance that I could’ve re-offended.”).

19. *Id.*

environment in which he was living at the Larned facility. His focus on self-improvement and strategies for what one might call pedophilia-sobriety seemed to arise only in the latter institution. Yet in Morris' approach, the punishment suffered over the years should have—if accomplished properly—conveyed to Hendricks that his behavior violated the surrounding society's norms and that he ought, therefore, to try to conform his behavior to those norms. The hospital setting, by contrast, could perhaps provide treatment for his behavioral pathology but would not convey to Hendricks, in an autonomy-regarding manner, the importance of the norms he had violated.

What in fact occurred, however, was nearly the opposite. When he was being punished, Hendricks showed no inclination toward self-awareness and rehabilitation. He simply wanted to do his time and get out, and he had no plans of stopping his offending conduct once he did. In the hospital, however, where he was viewed as a patient in need of treatment, he finally came to realize that what he had been doing was wrong and to take an active participatory role in attempting to modify his own patterns of behavior.

Morris could respond in a number of ways. First, our current prison system is hardly the model of individual-regarding retribution that would respectfully convey to an offender that he has violated the society's norms and thereby restore the fabric of his relationship with the outside world. As currently practiced, punishment disproportionately affects the poor and members of minority groups, and the substantive criminal law is itself extremely broad in ways that undermine the legitimacy of the entire system, by Morris' own estimation.²⁰ Second, Hendricks may be atypical in his responses, respectively, to punishment and civil commitment. And third, the "civil" confinement of Hendricks may well have been pretextual and thus, in fact, have constituted punishment of the humane sort that might

20. See Herbert Morris, *A Paternalistic Theory of Punishment*, *supra* note 1, at 269. Morris' theory of punishment:

includes at least two conditions that may be only marginally congruent with our social world. The first is that the norms addressed to persons are generally just and that the society is to some substantial extent one in which those who are liable to punishment have roughly equal opportunities to conform to those just norms. The second condition is equally important. The theory presupposes that there is a general commitment among persons to whom the norms apply to the values underlying them. If these two conditions are not met, we do not have what I understand as a practice of punishment for which any moral justification can be forthcoming.

Id.

succeed in reaching an offender.

Though these points have some weight, they do not persuade me that Hendricks' experience is either anomalous or represents an insignificant development for retributive theory. Our prison system is arbitrary in many ways and unfair to minorities and the poor, often punishing them for acts that should not have been criminal in the first place. Hendricks, however, is a white man who, though not wealthy, is financially comfortable (and in fact uses his own money to make purchases like the model airplane set he ultimately bought from a less judgmental supplier). His crime, moreover, is not the sort of offense that civil libertarians advocate should be legal—sexual abuse of children is not considered by many, beside active perpetrators, to be a victimless crime, like drug use or sexual acts between consenting adults. To the extent that the system is unfair, in other words, Hendricks himself did not experience that unfairness until he encountered the Kansas SVPA, which cheated him of the benefit of his plea bargain. And that was when he began to repent.

On the question of Hendricks' being atypical, the apparent factors distinguishing him from other criminals—his persistent recidivism and the vileness of his crime—would seem to make him an especially poor candidate for therapeutic insight. To the extent that he does perceive his stay in Larned as truly a form of punishment, he becomes defensive and unrepentant, speaking of never pleading guilty again, with the implicit threat that he might, in fact, *be* guilty again. Only when he turns his attention to therapists and caseworkers, those who seem focused on providing him with psychiatric help, does he willingly consider the evil of his past deeds, the pain he has caused to innocent children, and the work he intends to do to avoid offending again. By recognizing that something in him, the sexual desire for children, is a pathology, therapeutic intervention permits him to join forces with the professionals against the disease and simultaneously recognize the evil of his deeds.

Just as prison should not be held up as a model of punishment, the typical mental hospital for the criminally insane is no model for therapeutic intervention. As a practical matter, then, Morris' objections to the therapeutic state might be persuasive, even if, theoretically, they are not necessarily so. Like prisons, actual hospitals for the insane tend to reflect the contempt and lack of empathy that society actually feels for those who would violate its laws, whether or not through an exercise of true agency and responsibility. But just as Morris theorizes about

punishment as an ideal form rather than in its present reality, it is worth theorizing about the therapeutic state in its similarly ideal form, rather than as expressed in the ugly and abusive facilities that exist around the country.

Morris speaks of crime as creating a rift between the offender and society that must be mended through punishment. If one is interested in mending the rift, however, it is my contention that human nature and its response to different kinds of social control are a critical component of determining whether punishment is indeed the way to proceed. If the criminal perceives punishment as an adversarial event that signals victory and power over the vanquished, then the punishment will have no moral implications for the convict, any more than it would if dealt by a schoolyard bully. There is good reason to think that this is exactly how the criminal perceives punishment.

Consider recent popular writings in the area of conflict management, including such books as *Emotional Intelligence*, *Working With Emotional Intelligence*, *Difficult Conversations*, and *I Only Say This Because I Love You*.²¹ The message of all of these works is that conventional approaches to conflict management—seen as arguments in which people put each other in their places and insist on the rightness of their respective positions—serve only to escalate conflicts to a point at which no resolution is possible.

Each person becomes ego-invested to such an extent that he is unwilling to consider solutions that might be mutually beneficial. Underneath the bravado of each side's repeated assertions, oftentimes, is the strong sense that one's feelings and perspectives are being devalued, dismissed, and ignored by the "other side." That is the adversarial approach to conflict; fight it out, and let the better position win.

In some forms of alternative dispute resolution, including the informal methods of raising and resolving interpersonal problems proposed in the books mentioned above, the procedural objective is for people to feel like those around them are listening to and understanding their perspectives. Rather than simply imposing the more powerful person's preference on the weaker person, these approaches allow everyone to emerge with her ego intact. The contrast exemplifies the difference between punishment, in which the individual is ideally made

21. See DANIEL GOLEMAN, *EMOTIONAL INTELLIGENCE* (1995); DANIEL P. GOLEMAN, *WORKING WITH EMOTIONAL INTELLIGENCE* (1998); DOUGLAS STONE ET AL., *DIFFICULT CONVERSATIONS* (1999); DEBORAH TANNEN, *I ONLY SAY THIS BECAUSE I LOVE YOU* (2001).

to suffer as a means of conveying that he is wrong and society is right, and therapeutic intervention, in which the individual ideally comes to realize that something in his or her behavioral repertoire is harmful both to others and ultimately to himself. The therapeutic setting provides the means to address that harm, in the way an antibiotic might address a contagious disease. In other words, bypassing the individual's moral agency, ironically, may give that agency the space in which to form alliances with those attempting to treat the behavioral condition that gives rise to the offending conduct.

An important experiment in addressing low-level crime by alcoholics and drug addicts has taken the form of drug courts. Michael C. Dorf and Charles F. Sabel have written about these courts and the special challenges they pose to our conventional adversarial system.²² The drug courts allow a person who qualifies (because of an addiction, for example), who is charged with one of a set of relatively minor offenses, to opt into the drug court instead of attempting to defend himself within the conventional criminal justice system.²³ In drug court, the individual is prescribed a regimen of drug treatment, which includes periodic random tests, as well as the alternative of prison if the individual fails to cooperate adequately in his own recovery.

Because of the threat of imprisonment, drug treatment in this context is coercive rather than chosen. It is also a form of government-enforced therapy of the sort that would seem to violate the individual's dignity under Morris' analysis. From a defense attorney's perspective, moreover, the drug court structure might appear to give rise to a conflict of interest. This is because the attorney, instead of focusing single-mindedly on minimizing the amount of incarceration imposed upon his client, participates with the State in trying to figure out what would be "best" for the client, from both a personal and societal perspective. Perhaps confirming my general conclusions from the Hendricks example, though, drug courts appear to be much more effective than prison, both in reforming behavior,²⁴ and in reaching people.²⁵

22. Michael C. Dorf & Charles F. Sabel, *Drug Treatment Courts and Emergent Experimentalist Government*, 53 VAND. L. REV. 831 (2000).

23. *Id.* at 832 ("Persons charged with relatively low-level, non-violent criminal misconduct may opt out of the criminal trial court if the prosecuting attorney consents to the filing of charges in the treatment court. In treatment court, the defendant pleads guilty or otherwise accepts responsibility for a charged offense and accepts placement in a court-mandated program of drug treatment.").

24. Steven Belenko, *Research on Drug Courts: A Critical Review*, 1 NAT'L DRUG CT. INST. REV. 1 (1998).

Many theorists probably have an instinctively negative response to the whole idea of therapeutic approaches to people who have engaged in criminal misconduct. Behavioral treatments for “bad” people, for example, have acquired a negative reputation in the popular media. Films like *A Clockwork Orange*²⁶ suggest that such treatments are cruel and arbitrary. The nature of this reputation for cruelty, moreover, raises a second difficulty: perhaps operant conditioning is in fact a form of punishment. If so, might I be proposing that we substitute a cruel and arbitrary punishment for a more humane approach?

One could ultimately conclude that there is not much difference between behavioral treatment in a civil facility and criminal punishment in a prison. The U.S. Supreme Court has drawn an important constitutional line between the two, but some—including me—have criticized the line as artificial.²⁷ Consider behavioral aversion therapy.

*A Clockwork Orange*²⁸ critiques one kind of therapeutic approach to violence. Alex, the main character, a despicable person who savagely beats and rapes people for fun, is caught and held in a behavioral conditioning program in which he is given nausea-inducing drugs and then placed in front of a screen on which violent acts are depicted. Over time, he comes to have a visceral, nauseated reaction to violence, much as Pavlov’s dogs began to salivate at the sound of a dinner bell. Alex then returns to the world, where he is unable to act violently, because of his conditioning. Since the world outside has not changed, however, he is vulnerable to the ubiquitous violence that may have made him the way he used to be in the first place. In addition, because the scenes of violence exhibited during his incarceration were accompanied by background music by his favorite composer, Beethoven, he is also robbed of his taste for “Ludwig Van.” This is punishment indeed, and it changes its recipient into someone who literally *cannot* commit the

25. Even critics of drug courts acknowledge the benefits of the experience. See JAMES L. NOLAN, JR., *REINVENTING JUSTICE: THE AMERICAN DRUG COURT MOVEMENT* 194-204 (2001) (describing various drug court criticisms, and discussing justifications based upon the therapeutic benefits that drug courts offer the defendant).

26. *A CLOCKWORK ORANGE* (Warner Bros. 1971).

27. See, e.g., Sherry F. Colb, *The Three Faces of Evil*, 86 GEO. L.J. 677 (1998) (reviewing ELYN R. SAKS WITH STEPHEN H. BEHNKE, *JEKYLL ON TRIAL: MULTIPLE PERSONALITY DISORDER & CRIMINAL LAW* (1997)); Steven I. Friedland, *On Treatment, Punishment, and the Civil Commitment of Sex Offenders*, 70 U. COLO. L. REV. 73, 112 (1999) (“Not only is the Court’s conclusion about the civil nature of the Kansas law erroneous, but so is its analysis regarding treatment.”).

28. See *supra* note 26.

crimes of which he used to be guilty.²⁹

Though the aversion therapy in *A Clockwork Orange* is punishment, Morris would almost certainly not consider it a legitimate form of punishment.³⁰ It does deliberately inflict suffering upon the subject in response to his past misconduct, and thus would perhaps satisfy some retributivists. It does not, however, bring suffering as a component of reaching out and mending a rift that the target has created between himself and society. It does not address and communicate with him at an intellectual or moral level, but rather uses his biological systems as an instrument (or a weapon) for altering his behavior. The fact that Alex's love of Beethoven is demolished in the process perhaps exemplifies the extent to which the experience of aversion therapy bypasses the conscious, intellectual, and autonomous being.

In real life, some physicians have used forms of aversion therapy to try to "cure" people of homosexuality.³¹ Such cures have not proven successful, and the entire enterprise of "reparative therapy" is ethically questionable and independently troubling.³² However, if it did succeed,

29. *Id.*

30. See Morris, *A Paternalistic Theory*, *supra* note 1, at 269-70.

Excessively lengthy prison terms and the inhumane conditions under which they are served, for example, can be effectively criticized with a clear conception of the good defined by the paternalistic theory Punishments that are aimed at degrading or brutalizing a person are not conducive to moral awakening but only to bitterness and resentment The paternalistic position . . . implies that there is a non-waivable, non-forfeitable, non-relinquishable right—the right to one's status as a moral being, a right that is implied in one's being a possessor of any rights at all.

Id.

[U]nacceptable to this theory would be any response that sought the good of a wrongdoer in a manner that bypassed the human capacity for reflection, understanding, and revision of attitude that may result from such efforts. Any punitive response to a fully responsible being, then, and it might be no more than the giving of an evil-tasting pill or some form of conditioning, that directly in some causal way, with or without the agent's consent, sought to bring about a good, say instantaneous truth or aversion to acting violently, would be incompatible with this constraint.

Id. at 265-66.

31. Karen M. Jordan & Robert H. Deluty, *Clinical Interventions by Psychologists with Lesbians and Gay Men*, 51 J. CLINICAL PSYCHOL. 448 (1995); Paul Latimer, *A Case of Homosexuality Treated by In Vivo Desensitization and Assertive Training*, 22 CAN. PSYCHIATRIC ASS'N J. 185 (1977); Stephen J. Sansweet, *Aversion Therapy: Punishing of People to Change Behavior Gains Use, Controversy*, WALL STREET J., Jan. 2, 1974, at 1.

32. See Karolyn Ann Hicks, "Reparative" Therapy: Whether Parental Attempts to Change a Child's Sexual Orientation Can Legally Constitute Child Abuse, 49 AM. U. L. REV. 505, 510 (1999) ("[R]eparative' therapy [to change sexual orientation] on children

its mode of operation would be to bypass the conscious mind and manipulate people without appealing to their higher natures. In this regard, we inflict suffering on a target as the means to his alteration rather than as a communicative act designed to convey condemnation and address and reach the target's moral sensitivities.

If communication is the goal of punishment, however, one must consider what it means to communicate moral outrage effectively. Presumably, simply stating "you did a bad thing" would not suffice, or punishment would be unnecessary. Though Morris is not a utilitarian, his model for retribution does appear to presuppose—as all deontological theories do—some core principles about human nature. It assumes, for example, that one can communicate with another's higher mental faculties as a means of recommending attitude adjustment. One purpose of punishment—the infliction of suffering as a just response to crime—is to speak with emphasis. As mentioned earlier, it is akin to a metaphorical exclamation point at the end of the sentence: "you have done wrong and offended our way of life."

Social psychologists have convincingly demonstrated, however, that attitude adjustment ordinarily does not follow this pattern. When we wish to change attitudes, we begin by changing the way people behave. In other words, we proceed through manipulation. Their attitudes might then change in response to their actions.

III. LESSONS OF INTRINSIC MOTIVATION RESEARCH FOR RETRIBUTIVE THEORY

Over the last four decades, social psychologists have conducted important experiments on how to sustain or diminish people's "intrinsic motivation."³³ Intrinsic motivation refers to one's internal desire to

should . . . be considered abuse and neglect because of its damaging effects." "The generally accepted view among the medical and psychological professions is that efforts to alter a person's core gender identity are futile and unethical." Taylor Flynn, *Transforming the Debate: Why We Need to Include Transgender Rights in the Struggles for Sex and Sexual Orientation Equality*, 101 COLUM. L. REV. 392, 394-95 n.10 (2001) (citing GERALD MALLON, PRACTICE WITH TRANSGENDERED CHILDREN, IN SOCIAL SERVICES WITH TRANSGENDERED YOUTH 49, 55-58 (Gerrald Mallon ed., 1999)).

33. See, e.g., Rosemarie Anderson et al., *The Undermining and Enhancing of Intrinsic Motivation in Preschool Children* 34, J. PERSONALITY & SOC. PSYCHOL. 915 (1976); Edward L. Deci et al., *A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation*, 125 PSYCHOL. BULL. 627 (1999) (a comprehensive meta-analysis confirming the conclusions of the earlier work.); Mark R. Lepper et al., *Undermining Children's Intrinsic Interest with Extrinsic Reward: A Test of the 'Overjustification' Hypothesis*, J. PERSONALITY & SOC. PSYCHOL. 129 (1973). The

engage in particular activities. If I am intrinsically motivated to eat a cantaloupe, for example, I will eat it even if it serves no independent purpose beyond giving me pleasure. Most people are intrinsically motivated to eat and sleep but must be given extrinsic motivation to do schoolwork (as a child) or paid work (as an adult). The psychological studies I reference explore the question of how people come to lose intrinsic motivation and whether, in particular, such loss might be externally manipulated.

The study designs involved an experimenter providing subjects with a goal-oriented activity in which to engage. After assessing the subjects' intrinsic interest in the activity, experimenters would tell some subjects that they would receive an award for engaging in the activity again and would tell other subjects nothing about awards. Subjects would then be given a second opportunity to engage in the activity. Those promised awards would receive the awards, and half of those not promised awards would receive surprise awards. Finally, on a third occasion, the subjects would be given an opportunity to engage in the activity without any possibility of an award.

The results showed that subjects who had been told in advance that they would receive an award in the second round of the experiment tended to lose some of their initial interest in the activity by the third opportunity, when no external award was given. The subjects who had *not* been told about an award in advance, by contrast, maintained their initial interest in the activity, whether or not they later received a surprise award.³⁴ The study demonstrated that simply anticipating a reward for an activity previously enjoyed can reduce a participant's intrinsic motivation to engage in that activity in the future.

One variant on this set of experiments involved children, ages three to five years old, who were provided with paper and markers on three

overjustification hypothesis, however, has engendered heated controversy among behaviorists. See, e.g., Judy Cameron & W. David Pierce, *Reinforcement, Reward, and Intrinsic Motivation: A Meta-Analysis*, REV. OF EDUC. RES. 363 (1994) ("A recent, comprehensive meta-analysis (Deci, Koestner, & Ryan, 1999) confirmed, in spite of claims to the contrary . . . that all expected tangible rewards made contingent on task performance do reliably undermine intrinsic motivation."); Robert Eisenberger & Judy Cameron, *Detrimental Effects of Reward: Reality or Myth*, 51 AM. PSYCHOL. 1153 (1996); Richard M. Ryan & Edward L. Deci, *Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being*, 55 AM. PSYCH. 68 (2000).

34. See generally Richard M. Ryan & Edward L. Deci, *Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being*, 55 AM. PSYCH. 68 (2000).

occasions. During the second exposure to the paper and markers, children were subject to one of the three experimental manipulations: expected reward, surprise reward, or no reward.³⁵ The children in the expected reward condition experienced diminished intrinsic motivation to play with the markers when no reward was available.

The theory that psychologists such as Edward L. Deci, Richard M. Ryan, Mark R. Lepper, David Greene, and Richard E. Nisbett developed to explain such results is the “overjustification hypothesis.”³⁶ The theory holds that when a person engages in an activity, she watches herself in the same way as she would watch other people engaging in an activity. She then draws inferences about her likes and dislikes from her observations, and these inferences can either strengthen, diminish, or reinforce her own attitudes about what she likes and dislikes. If her behavior is “overjustified” by external factors, then she will attribute the behavior to those factors rather than to her own intrinsic interest.

For example, if I see my husband, a law professor, read a romance novel, I will infer that he enjoys reading such novels. I might, therefore, decide to buy him one for his birthday. If, on the other hand, I see my husband reading a law review article, I will not draw a parallel inference and buy him a new law review article for his birthday. This is because I know that such readings are part of his job and this fact fully accounts for his activity without resort to any attribution of preference. Similarly, if I am the one reading the law review article in my field, I will probably not infer that I especially enjoy reading such material. This is true even if I might have actually enjoyed reading such articles in the absence of the work incentive.

To take a more mundane example, if I see a man spontaneously order pizza with mushrooms at an Italian restaurant, I will infer that he likes pizza and that he also likes mushrooms. If, on the other hand, I see him attempt to order several other types of pizza (plain, Sicilian, and peppers and onions, for example) and only order mushroom when he is told the other types will not be available for an hour, I will probably infer that he likes pizza, does not necessarily like mushrooms very much but is hungry enough to eat them. Similarly, if I am the one doing the ordering, I might infer that I am not such a big fan of mushrooms. It may be that I previously liked mushrooms, but now that I have acted as though I do not, observing this action could influence my attitude toward mushrooms in the future.

35. See Lepper et al., *supra* note 33.

36. *Id.*

When it comes to food, of course, one probably has relatively stable preferences and, therefore, need not rely primarily on self-observation to determine what one likes. For other activities, however, such as reading, one's preordained preferences might not be as clear. The child who draws a picture after learning that he will receive an external reward for doing so could infer later that the only reason he drew the picture was to get the reward. Positive feelings associated with the drawing activity might, accordingly, be attributed to the expected reward.

Later, when no reward is available, the child might choose not to draw. He understands himself to have no intrinsic preference for drawing and he is, therefore, not inclined to draw. If, on the other hand, the child draws while expecting no external reward, then the memory of the activity is "I chose to draw pictures because I felt like it." He is accordingly more likely to want to draw again. Therefore, when we communicate directly to children or adults, "we will give you a reward for doing something that is a good thing to do," the response might be to think, "I don't really like doing that activity and will only do it when I can once again obtain a reward."

This phenomenon of diminishing interest in something as a response to being explicitly rewarded for it has important implications for the functioning of punishment as a mode of communication.³⁷ Consider the man who molests a child. He is tried, convicted, and placed in a prison for ten years as his punishment. The State's goal, on Morris' analysis, is—or ought to be—to emphasize to the man that his behavior was wrong, that it violated significant social norms embraced by his community. While suffering his imprisonment, however, if the man operates in the way the subjects in the above experiments did, he might instead draw a very different set of inferences.

First, the man may conclude that he molested children because he loves molesting children, and he must have enjoyed it a great deal to have done it in spite of the punishment he was warned about and now suffers. Once he gets out, he thinks, he might try to do it again, especially if he can get away with it. Though the government believes what he did was wrong, he might conclude that he disagrees with the

37. Indeed, some of the studies suggested as much. See Ryan & Deci, *supra* note 34, at 70 ("[R]esearch revealed that not only tangible rewards but also threats, deadlines, directives, pressured evaluations, and imposed goals diminish intrinsic motivation because, like tangible rewards, they conduce toward an external perceived locus of causality.").

government. Whatever happens to him, he can thereby maintain the belief that what he did is fine.

If this hypothesis holds, then rather than mend the rift between him and the surrounding community, the infliction of punishment as deliberate condemnation or as the message “you’re wrong!” with an exclamation point is likely to result in the hardening of the criminal’s attitudes and the concomitant response of “no, I’m not!” This appears to be what repeatedly happened to Leroy Hendricks as he served multiple prison sentences but failed to turn away from—or even to reflect on the wrongfulness of—his illegal actions. “I must really like and believe in child molestation,” he may have thought, “because the state’s infliction of suffering has not had any impact on what I do.” Just as reward may undermine intrinsic motivation, punishment may enhance it.

By the time I interviewed him, Hendricks had begun to view his actions as wrong. This occurred as a result of a number of phenomena, each of which seems at odds with Morris’ prescription for treating a person with dignity and respect through punishment, rather than as a subject to be manipulated and thus transformed through therapeutic intervention. The first phenomenon was the understanding of his desire to have sexual relations with children as a pathology from which he suffers, one which the doctors working at the Larned facility wish to help him overcome.

Rather than demanding that he view his prior bad acts as evil, the structure of his confinement conveys the view that he is a victim of something, and he and the mental health professionals can battle this something together. He then receives a form of therapy that asks him to take the perspective of the people his behavior has harmed. The message is, “see how your condition, when untreated, makes others suffer.” His prior behavior is posited as somehow external to him, and he can then reject it without simultaneously rejecting the person he once was.

Ironically, once he is classified as having been—in a sense—unable to control himself, he can more easily begin to exercise autonomy and plan strategies for altering his behavior patterns. He undergoes a sincere change of heart, of the sort that Morris might expect from someone receiving honest, respectful retribution, and he simultaneously forms an alliance with those who wish to protect the people he would victimize.

Hendricks continues, interestingly, to reject punishment and

despises Attorney General Carla Stovall of Kansas for wanting to punish him after his time for punishment has passed. It is precisely through a bypass of his moral agency as the source of his wrongdoing that he comes, for the first time, to feel remorse for that wrongdoing. He can now feel remorse without the accompanying loss of face that accompanies apology during condemnation.

Consider as illustrative the aversion to apologizing. When two friends have had an unpleasant argument, they often wish to be friends again and yet have a difficult time resolving their differences. Each one wants the other to apologize, and neither is willing to be the one to do it. If an apology is eventually to come, moreover, it is *least* likely to happen while one is condemning the other. Instead, as one says "what you did is wrong. You owe me an apology," the other feels or even says, "I have done nothing wrong. I have nothing to apologize for. You're the one who owes me an apology." Each combatant may mobilize arguments in defense of his respective position, while the other disregards these arguments in his efforts to come up with more of his own. This is the adversarial system at work, and each person is accordingly motivated to rebut what the other says and to resist the implications of his interlocutor's position. Though it may provide a useful way for a third party to decide who is right, as in an adversarial court proceeding, this is no recipe for bilateral reconciliation.

If, as Morris suggests, one reason for punishment is to permit reconciliation between the offender and the community, a kind of reintegration into society, then, as I have argued here, punishment could be inherently bound to fail. It may be that punishment communicates society's outrage, in the sense of informing the punished individual that society disapproves of what he has done, but this communication is bound to anger him and make him defensive rather than create a space in which he can truly repent.

My analysis proceeds on the assumption that mending the rift between society and wrongdoer through punishment cannot be said to take place if the predictable, real-world impact of condemnation through punishment is to pressure the wrongdoer, perhaps unconsciously, toward a stronger embrace of the disapproved norm. Turn again to the example of the wayward child.

If a child becomes angry and belligerent when punished, then how do parents train their children to behave? Some would say that parents do not train their children to behave and that children simply imitate the

behavior of their same-age peers.³⁸ But this is surely not the whole story, because peers do not ordinarily speak to every aspect of a child's life. Parents shape at least some of the behavior of their children, for better or for worse, and it is useful to examine the different methods by which they do so.

I have had a human child for only a short time now, but I have had two dogs for many years. I often find myself learning about behavior modification in ways that are surprisingly relevant to human interaction. Take punishment, for example. I do not personally believe in corporal punishment and therefore do not ordinarily discipline my dogs by hitting them. I do, however, say "no" and "bad" in a stern voice, which they do not like. One thing I taught them with this "no" method was that they are not allowed to lie on the living room couch. It seemed to have been a big success, until one day, I came home earlier than usual to catch them unawares, lying on the couch.

It became clear that they learned I would punish them only if they were on the couch in my presence, not that they *should not* go onto the couch at all. When they could evade punishment, then, going on the couch provoked no guilt feelings in them. My reward methods were similarly unsuccessful. I gave biscuits to the older one whenever he listened to me (e.g., "come," "sit," etc.). I discovered soon, however, that when I told him to do things, he looked to see if I had biscuits, and if I didn't, he proceeded to ignore the commands.

Much more successful, I found, was my and my husband's efforts at training the two dogs to remain within the periphery of an area in the park where we took them to play. Though they enjoyed running off to prohibited areas—and occasionally still do it—they have mostly stopped. We played ball with them daily and kept the ball and the running confined to the area where dogs are permitted to go. As a result, they "watched" themselves running in that area and have had an opportunity to conclude that this is where they like to play. Now, when we pass this area without stopping, they cry and try to pull us toward it.

We have manipulated them into loving this random area of the park, not by punishing or rewarding, but by repeatedly playing with them there. They did not run off—though they "could have"—because they were distracted. But from an observer's view—and they too observed their own behavior—it became plausible to conclude that they chose to stay where they did because they preferred that area to the rest

38. See JUDITH RICH HARRIS, *THE NURTURE ASSUMPTION: WHY CHILDREN TURN OUT THE WAY THEY DO* (1998).

of the park.

Just as we may have manipulated our dogs into thinking they love a small area in the park, the therapists treating Leroy Hendricks might be manipulating him by saying “you did these acts to children because you were operating under a false set of beliefs.” The truth may be that Hendricks acted on his sexual impulses because he did not care about the children and chose to place his own gratification higher on his priority list than the interests of his helpless victims.

By providing an alternative plausible story, however, even one that might be less descriptively accurate, the therapists permit Hendricks to see his earlier offending as “overjustified,” in the way the children who drew pictures for rewards saw their drawing in the intrinsic motivation studies. By conceptualizing his earlier behavior in this way, he can experience the decision to stop harming children as authentically his own. Through therapeutic manipulation, then, he acquires a level of moral agency that he lacked before.

In expressing enthusiasm for the civil confinement of some dangerous criminals, I am cognizant of the risk that the therapeutic state can turn into a warehouse for outcasts rather than an honest place of rehabilitative behavior modification. Perhaps that was Kansas’ true purpose, for example, in passing the SVPA and in confining Leroy Hendricks in Larned. But the incarceration punishment alternative is also risky. It has been quite unsuccessful in “reaching” those within its walls.

I propose here that, as in the experience of Leroy Hendricks, the therapeutic approach to wrongdoing—as a supplement or follow up to punishment—may be theoretically consistent with the possibility of repentance because of its structure, while the institution of punishment may not be. That in itself could be a feature of human, and perhaps all social animals’ nature. I do not suggest that punishment should therefore be abolished. There are good reasons to punish guilty people even if, perhaps *especially* if, those people are defiant and unrepentant in the face of society’s assertion of contrary norms. These reasons include desert, a society’s basic retributive interest, along with the hope of deterring others who can observe punishment and learn to avoid crime without having to lose face in the way that an offender does when he goes through the punishment himself. To reach and change the offender who is punished, however, therapeutic intervention may be necessary as well. For that reason, post-prison confinement like that of Leroy Hendricks in Larned, Kansas, could be an essential part of a

criminal justice scheme in which society is ultimately able to convey its norms effectively to those who repeatedly violate them.